2 The Participants

A brief introduction to the groups collaborating on the three projects:

Computer Science Projects

Globus Project (*www.globus.org*): fundamental technologies needed to build computational grids. The open source, open architecture Globus Toolkit provides the foundation for most grids worldwide.

Condor Project (www.cs.wisc.edu/condor): mechanisms and policies for supporting High Throughput Computing (HTC) on large collections of distributively owned computing resources.

Storage Resource Manager Collaboratory (*sdm.lbl.gov*): managing access to large amounts of data distributed over the grid. SRMs may be associated with each storage resource on the grid.

Storage Resource Broker (*www.sdsc.edu/dice/SRB*): providing a uniform interface over a network and accessing replicated data sets.

Other collaborating institutions: University of California at Berkeley, University of Southern California at San Diego, Northwestern University and Indiana University.

Experiments

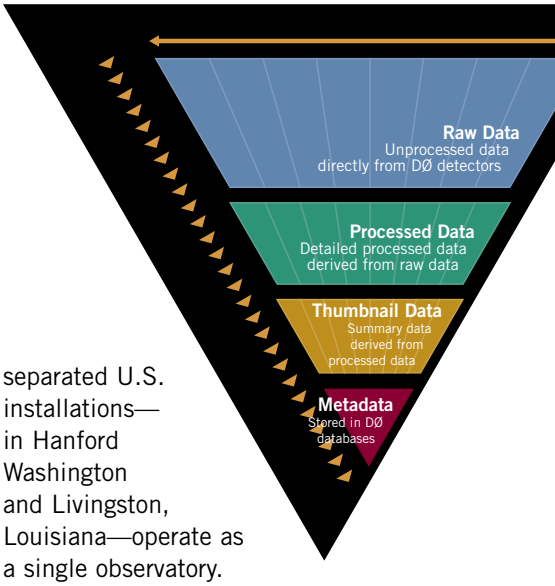
ATLAS (*www.usatlas.bnl.gov*), under construction for CERN's Large Hadron Collider (together with CMS): exploring the fundamental nature of matter and the basic forces that shape our universe, including the discovery of the Higgs particle.

BaBar (www.slac.stanford.edu/BFROOT): studying millions of B mesons produced by the PEP-II storage ring at SLAC. The collaboration consists of approximately 600 physicists and engineers from 76 institutions in 9 countries.

Compact Muon Solenoid (*uscms.fnal.gov*): exploring high transverse momentum physics at the LHC. Major goal (with ATLAS): discovering the Higgs particle, predicted by current theories but yet to be observed experimentally.

DØ (*www-DØ.fnal.gov*): precise studies of proton-antiproton interactions at the highest available energies at Fermi National Accelerator Laboratory's Tevatron Collider.

Laser Interferometer Gravitational-Wave Observatory (LIGO) (www.ligo.org): detecting cosmic gravitational waves and harnessing them for scientific research. Two widely separated U.S. installations—in Hanford Washington and Livingston, Louisiana—operate as a single observatory.



The National Virtual Observatory (NVO) (*www.us-vo.org*): uniting astronomy databases of many earthbound and orbital observatories. Goal: making the astronomy data from many surveys accessible to professional researchers, amateur astronomers, and students.

Sloan Digital Sky Survey (SDSS) (*www.sdss.org*): mapping in detail one-quarter of the entire sky; determining positions and absolute brightness of more than 100 million celestial objects; measuring distances to more than a million galaxies and quasars; examining large-scale patterns of galactic sheets and voids in the universe.

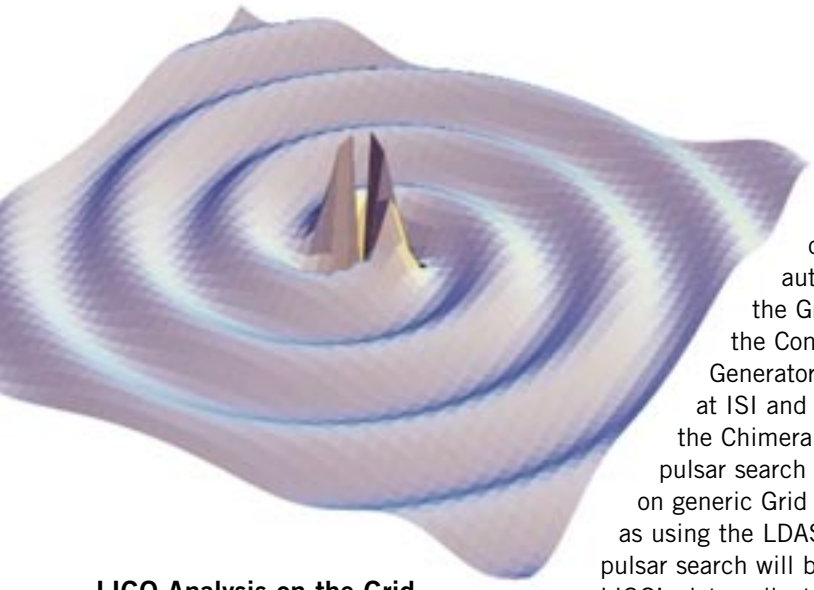
STAR (*www.star.bnl.gov*): searching for signatures of quark-gluon plasma (QGP) formation, and investigating the behavior of strongly-interacting matter at high-energy density at the Brookhaven National Lab's RHIC accelerator.

Thomas Jefferson National Accelerator Facility (TJNAF) (*www.jlab.org*): experiments using a high-intensity, high-energy, continuous electron beam to study the quark substructure of nucleons and nuclei.

Outreach and Education

University of Texas-Brownsville: leading community education and outreach initiatives, ensuring our work is understood and can be adapted by others.

Collaborations in Action



LIGO Analysis on the Grid

The Laser Interferometer Gravitational Wave Observatory (LIGO) and the GriPhyN project have been collaborating on a pulsar search application. The data involved is a long stretch (~4 months, $2x10^{11}$ points) of a single channel: the gravitational wave strain channel. The power spectra of the small segments are stacked to make a large frequency-time image, perhaps $4x10^{5}$ on each side. The pulsar search seeks coherent signals in this image. A source would appear on the frequency-time image as a wavering line, whose frequency might be 1 kHz but would be modulated by a few parts in 106 over periods of 1 day, and by a few parts in 104 over periods of 1 year. The data will also encode any secular variations of the source due to slowing of its rotational period.

The pulsar search requires greater computing and data resources than those available in the LIGO Scientific Collaboration. GriPhyN has used the Chimera Virtual Data Language to represent the pulsar search as an abstract workflow, which can be mapped automatically onto the Grid resources by the Concrete Workflow Generator (CWG) developed at ISI and integrated into the Chimera system. The pulsar search analysis can run on generic Grid resources as well as using the LDAS system. The pulsar search will be conducted using LIGO's data collected during the first scientific run of the instrument and will target a set of 1,000 known pulsar locations as well as random locations in the sky. The results of the analysis will be published via LDAS to the LIGO Scientific Collaboration.

DØ JIM

Job Dispatch Based on Availability of Needed Data Files

The DØ detector is collecting data at about 15 Mbytes/sec at the Tevatron—a rate greater than Fermilab's ability to provide sufficient compute cycles, to offer 500 physicists the timely feedback they need on the quality of the data and the physics results. DØ developed a distributed data handling and distributed data caching system (SAM), an automated support offering efficient delivery of data to offsite institutions on request. This system has been extended to include the automated scheduling of user data processing "jobs" over a computational Grid, with planning and policy applied to determine the best location for the job to execute, based on the availability of the needed data. The Job and Information Management (JIM) extensions to SAM will increase the total throughput and efficiency of data analysis without increasing the support and administrative load.

DØ is collaborating closely with the computer science groups to extend functionality and provide production quality features to the software. Remote job execution and monitoring is provided by the GRAM protocol and jobmanager service from the Globus package. The monitoring framework is provided by the Globus MDS system. Condor-G software provides reliable job submission and a system wide job manager, and the Condor Match Making system is used as the request broker.